

RECEIVED

MAY 05 2003

TECH CENTER 1600/2500

1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/009,383

DATE: 04/29/2003

TIME: 08:07:59

Input Set : A:\07763-043001.TXT

Output Set: N:\CRF4\04292003\J009383.raw

4 <110> APPLICANT: Gennaro, Maria L.
 6 <120> TITLE OF INVENTION: PROTEINS EXPRESSED BY MYCOBACTERIUM
 7 TUBERCULOSIS AND NOT BY BCG AND THEIR USE AS DIAGNOSTIC
 8 REAGENTS AND VACCINES
 10 <130> FILE REFERENCE: 07763-043001
 12 <140> CURRENT APPLICATION NUMBER: 10/009,383
 13 <141> CURRENT FILING DATE: 2001-11-02
 15 <150> PRIOR APPLICATION NUMBER: PCT/US00/12257
 16 <151> PRIOR FILING DATE: 2000-05-04
 18 <150> PRIOR APPLICATION NUMBER: 60/132,505
 19 <151> PRIOR FILING DATE: 1999-05-04
 21 <160> NUMBER OF SEQ ID NOS: 16
 23 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 591
 27 <212> TYPE: PRT
 28 <213> ORGANISM: Mycobacterium tuberculosis
 30 <400> SEQUENCE: 1
 31 Met Thr Ala Glu Pro Glu Val Arg Thr Leu Arg Glu Val Val Leu Asp
 32 1 5 10 15
 33 Gln Leu Gly Thr Ala Glu Ser Arg Ala Tyr Lys Met Trp Leu Pro Pro
 34 20 25 30
 35 Leu Thr Asn Pro Val Pro Leu Asn Glu Leu Ile Ala Arg Asp Arg Arg
 36 35 40 45
 37 Gln Pro Leu Arg Phe Ala Leu Gly Ile Met Asp Glu Pro Arg Arg His
 38 50 55 60
 39 Leu Gln Asp Val Trp Gly Val Asp Val Ser Gly Ala Gly Gly Asn Ile
 40 65 70 75 80
 41 Gly Ile Gly Gly Ala Pro Gln Thr Gly Lys Ser Thr Leu Leu Gln Thr
 42 85 90 95
 43 Met Val Met Ser Ala Ala Ala Thr His Ser Pro Arg Asn Val Gln Phe
 44 100 105 110
 45 Tyr Cys Ile Asp Leu Gly Gly Gly Leu Ile Tyr Leu Glu Asn Leu
 46 115 120 125
 47 Pro His Val Gly Gly Val Ala Asn Arg Ser Glu Pro Asp Lys Val Asn
 48 130 135 140
 49 Arg Val Val Ala Glu Met Gln Ala Val Met Arg Gln Arg Glu Thr Thr
 50 145 150 155 160
 51 Phe Lys Glu His Arg Val Gly Ser Ile Gly Met Tyr Arg Gln Leu Arg
 52 165 170 175
 53 Asp Asp Pro Ser Gln Pro Val Ala Ser Asp Pro Tyr Gly Asp Val Phe
 54 180 185 190
 55 Leu Ile Ile Asp Gly Trp Pro Gly Phe Val Gly Glu Phe Pro Asp Leu

ENTERED

RAW SEQUENCE LISTING

DATE: 04/29/2003

PATENT APPLICATION: US/10/009,383

TIME: 08:07:59

Input Set : A:\07763-043001.TXT

Output Set: N:\CRF4\04292003\J009383.raw

```

56          195          200          205
57 Glu Gly Gln Val Gln Asp Leu Ala Ala Gln Gly Leu Ala Phe Gly Val
58          210          215          220
59 His Val Ile Ile Ser Thr Pro Arg Trp Thr Glu Leu Lys Ser Arg Val
60 225          230          235          240
61 Arg Asp Tyr Leu Gly Thr Lys Ile Glu Phe Arg Leu Gly Asp Val Asn
62          245          250          255
63 Glu Thr Gln Ile Asp Arg Ile Thr Arg Glu Ile Pro Ala Asn Arg Pro
64          260          265          270
65 Gly Arg Ala Val Ser Met Glu Lys His His Leu Met Ile Gly Val Pro
66          275          280          285
67 Arg Phe Asp Gly Val His Ser Ala Asp Asn Leu Val Glu Ala Ile Thr
68          290          295          300
69 Ala Gly Val Thr Gln Ile Ala Ser Gln His Thr Glu Gln Ala Pro Pro
70 305          310          315          320
71 Val Arg Val Leu Pro Glu Arg Ile His Leu His Glu Leu Asp Pro Asn
72          325          330          335
73 Pro Pro Gly Pro Glu Ser Asp Tyr Arg Thr Arg Trp Glu Ile Pro Ile
74          340          345          350
75 Gly Leu Arg Glu Thr Asp Leu Thr Pro Ala His Cys His Met His Thr
76          355          360          365
77 Asn Pro His Leu Leu Ile Phe Gly Ala Ala Lys Ser Gly Lys Thr Thr
78          370          375          380
79 Ile Ala His Ala Ile Ala Arg Ala Ile Cys Ala Arg Asn Ser Pro Gln
80 385          390          395          400
81 Gln Val Arg Phe Met Leu Ala Asp Tyr Arg Ser Gly Leu Leu Asp Ala
82          405          410          415
83 Val Pro Asp Thr His Leu Leu Gly Ala Gly Ala Ile Asn Arg Asn Ser
84          420          425          430
85 Ala Ser Leu Asp Glu Ala Val Gln Ala Leu Ala Val Asn Leu Lys Lys
86          435          440          445
87 Arg Leu Pro Pro Thr Asp Leu Thr Thr Ala Gln Leu Arg Ser Arg Ser
88          450          455          460
89 Trp Trp Ser Gly Phe Asp Val Val Leu Leu Val Asp Asp Trp His Met
90 465          470          475          480
91 Ile Val Gly Ala Ala Gly Gly Met Pro Pro Met Ala Pro Leu Ala Pro
92          485          490          495
93 Leu Leu Pro Ala Ala Ala Asp Ile Gly Leu His Ile Ile Val Thr Cys
94          500          505          510
95 Gln Met Ser Gln Ala Tyr Lys Ala Thr Met Asp Lys Phe Val Gly Ala
96          515          520          525
97 Ala Phe Gly Ser Gly Ala Pro Thr Met Phe Leu Ser Gly Glu Lys Gln
98          530          535          540
99 Glu Phe Pro Ser Ser Glu Phe Lys Val Lys Arg Arg Pro Pro Gly Gln
100 545          550          555          560
101 Ala Phe Leu Val Ser Pro Asp Gly Lys Glu Val Ile Gln Ala Pro Tyr
102          565          570          575
103 Ile Glu Pro Pro Glu Glu Val Phe Ala Ala Pro Pro Ser Ala Gly
104          580          585          590

```

RAW SEQUENCE LISTING

DATE: 04/29/2003

PATENT APPLICATION: US/10/009,383

TIME: 08:07:59

Input Set : A:\07763-043001.TXT

Output Set: N:\CRF4\04292003\J009383.raw

```

106 <210> SEQ ID NO: 2
107 <211> LENGTH: 99
108 <212> TYPE: PRT
109 <213> ORGANISM: Mycobacterium tuberculosis
111 <400> SEQUENCE: 2
112 Met Glu Lys Met Ser His Asp Pro Ile Ala Ala Asp Ile Gly Thr Gln
113 1 5 10 15
114 Val Ser Asp Asn Ala Leu His Gly Val Thr Ala Gly Ser Thr Ala Leu
115 20 25 30
116 Thr Ser Val Thr Gly Leu Val Pro Ala Gly Ala Asp Glu Val Ser Ala
117 35 40 45
118 Gln Ala Ala Thr Ala Phe Thr Ser Glu Gly Ile Gln Leu Leu Ala Ser
119 50 55 60
120 Asn Ala Ser Ala Gln Asp Gln Leu His Arg Ala Gly Glu Ala Val Gln
121 65 70 75 80
122 Asp Val Ala Arg Thr Tyr Ser Gln Ile Asp Asp Gly Ala Ala Gly Val
123 85 90 95
124 Phe Ala Glu
127 <210> SEQ ID NO: 3
128 <211> LENGTH: 368
129 <212> TYPE: PRT
130 <213> ORGANISM: Mycobacterium tuberculosis
132 <400> SEQUENCE: 3
133 Met Leu Trp His Ala Met Pro Pro Glu Leu Asn Thr Ala Arg Leu Met
134 1 5 10 15
135 Ala Gly Ala Gly Pro Ala Pro Met Leu Ala Ala Ala Ala Gly Trp Gln
136 20 25 30
137 Thr Leu Ser Ala Ala Leu Asp Ala Gln Ala Val Glu Leu Thr Ala Arg
138 35 40 45
139 Leu Asn Ser Leu Gly Glu Ala Trp Thr Gly Gly Gly Ser Asp Lys Ala
140 50 55 60
141 Leu Ala Ala Ala Thr Pro Met Val Val Trp Leu Gln Thr Ala Ser Thr
142 65 70 75 80
143 Gln Ala Lys Thr Arg Ala Met Gln Ala Thr Ala Gln Ala Ala Ala Tyr
144 85 90 95
145 Thr Gln Ala Met Ala Thr Thr Pro Ser Leu Pro Glu Ile Ala Ala Asn
146 100 105 110
147 His Ile Thr Gln Ala Val Leu Thr Ala Thr Asn Phe Phe Gly Ile Asn
148 115 120 125
149 Thr Ile Pro Ile Ala Leu Thr Glu Met Asp Tyr Phe Ile Arg Met Trp
150 130 135 140
151 Asn Gln Ala Ala Leu Ala Met Glu Val Tyr Gln Ala Glu Thr Ala Val
152 145 150 155 160
153 Asn Thr Leu Phe Glu Lys Leu Glu Pro Met Ala Ser Ile Leu Asp Pro
154 165 170 175
155 Gly Ala Ser Gln Ser Thr Thr Asn Pro Ile Phe Gly Met Pro Ser Pro
156 180 185 190
157 Gly Ser Ser Thr Pro Val Gly Gln Leu Pro Pro Ala Ala Thr Gln Thr
158 195 200 205

```

RAW SEQUENCE LISTING

DATE: 04/29/2003

PATENT APPLICATION: US/10/009,383

TIME: 08:07:59

Input Set : A:\07763-043001.TXT

Output Set: N:\CRF4\04292003\J009383.raw

```

159 Leu Gly Gln Leu Gly Glu Met Ser Gly Pro Met Gln Gln Leu Thr Gln
160      210                      215                      220
161 Pro Leu Gln Gln Val Thr Ser Leu Phe Ser Gln Val Gly Gly Thr Gly
162 225                      230                      235                      240
163 Gly Gly Asn Pro Ala Asp Glu Glu Ala Ala Gln Met Gly Leu Leu Gly
164                      245                      250                      255
165 Thr Ser Pro Leu Ser Asn His Pro Leu Ala Gly Gly Ser Gly Pro Ser
166                      260                      265                      270
167 Ala Gly Ala Gly Leu Leu Arg Ala Glu Ser Leu Pro Gly Ala Gly Gly
168                      275                      280                      285
169 Ser Leu Thr Arg Thr Pro Leu Met Ser Gln Leu Ile Glu Lys Pro Val
170      290                      295                      300
171 Ala Pro Ser Val Met Pro Ala Ala Ala Ala Gly Ser Ser Ala Thr Gly
172 305                      310                      315                      320
173 Gly Ala Ala Pro Val Gly Ala Gly Ala Met Gly Gln Gly Ala Gln Ser
174                      325                      330                      335
175 Gly Gly Ser Thr Arg Pro Gly Leu Val Ala Pro Ala Pro Leu Ala Gln
176                      340                      345                      350
177 Glu Arg Glu Glu Asp Asp Glu Asp Asp Trp Asp Glu Glu Asp Asp Trp
178      355                      360                      365
180 <210> SEQ ID NO: 4
181 <211> LENGTH: 100
182 <212> TYPE: PRT
183 <213> ORGANISM: Mycobacterium tuberculosis
185 <400> SEQUENCE: 4
186 Met Ala Glu Met Lys Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly
187 1      5      10      15
188 Asn Phe Glu Arg Ile Ser Gly Asp Leu Lys Thr Gln Ile Asp Gln Val
189      20      25      30
190 Glu Ser Thr Ala Gly Ser Leu Gln Gly Gln Trp Arg Gly Ala Ala Gly
191      35      40      45
192 Thr Ala Ala Gln Ala Ala Val Val Arg Phe Gln Glu Ala Ala Asn Lys
193      50      55      60
194 Gln Lys Gln Glu Leu Asp Glu Ile Ser Thr Asn Ile Arg Gln Ala Gly
195 65      70      75      80
196 Val Gln Tyr Ser Arg Ala Asp Glu Glu Gln Gln Ala Leu Ser Ser
197      85      90      95
198 Gln Met Gly Phe
199      100
201 <210> SEQ ID NO: 5
202 <211> LENGTH: 666
203 <212> TYPE: PRT
204 <213> ORGANISM: Mycobacterium tuberculosis
206 <400> SEQUENCE: 5
207 Met Ala Ala Asp Tyr Asp Lys Leu Phe Arg Pro His Glu Gly Met Glu
208 1      5      10      15
209 Ala Pro Asp Asp Met Ala Ala Gln Pro Phe Phe Asp Pro Ser Ala Ser
210      20      25      30
211 Phe Pro Pro Ala Pro Ala Ser Ala Asn Leu Pro Lys Pro Asn Gly Gln

```

RAW SEQUENCE LISTING

DATE: 04/29/2003

PATENT APPLICATION: US/10/009,383

TIME: 08:07:59

Input Set : A:\07763-043001.TXT

Output Set: N:\CRF4\04292003\J009383.raw

212		35		40		45											
213	Thr	Pro	Pro	Pro	Thr	Ser	Asp	Asp	Leu	Ser	Glu	Arg	Phe	Val	Ser	Ala	
214		50					55					60					
215	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Thr	Pro	Met	
216	65					70					75					80	
217	Pro	Ile	Ala	Ala	Gly	Glu	Pro	Pro	Ser	Pro	Glu	Pro	Ala	Ala	Ser	Lys	
218					85					90					95		
219	Pro	Pro	Thr	Pro	Pro	Met	Pro	Ile	Ala	Gly	Pro	Glu	Pro	Ala	Pro	Pro	
220				100					105					110			
221	Lys	Pro	Pro	Thr	Pro	Pro	Met	Pro	Ile	Ala	Gly	Pro	Glu	Pro	Ala	Pro	
222			115				120					125					
223	Pro	Lys	Pro	Pro	Thr	Pro	Pro	Met	Pro	Ile	Ala	Gly	Pro	Ala	Pro	Thr	
224		130					135					140					
225	Pro	Thr	Glu	Ser	Gln	Leu	Ala	Pro	Pro	Arg	Pro	Pro	Thr	Pro	Gln	Thr	
226	145					150					155					160	
227	Pro	Thr	Gly	Ala	Pro	Gln	Gln	Pro	Glu	Ser	Pro	Ala	Pro	His	Val	Pro	
228					165					170					175		
229	Ser	His	Gly	Pro	His	Gln	Pro	Arg	Arg	Thr	Ala	Pro	Ala	Pro	Pro	Trp	
230			180						185					190			
231	Ala	Lys	Met	Pro	Ile	Gly	Glu	Pro	Pro	Pro	Ala	Pro	Ser	Arg	Pro	Ser	
232			195					200				205					
233	Ala	Ser	Pro	Ala	Glu	Pro	Pro	Thr	Arg	Pro	Ala	Pro	Gln	His	Ser	Arg	
234		210					215					220					
235	Arg	Ala	Arg	Arg	Gly	His	Arg	Tyr	Arg	Thr	Asp	Thr	Glu	Arg	Asn	Val	
236	225					230					235					240	
237	Gly	Lys	Val	Ala	Thr	Gly	Pro	Ser	Ile	Gln	Ala	Arg	Leu	Arg	Ala	Glu	
238					245					250					255		
239	Glu	Ala	Ser	Gly	Ala	Gln	Leu	Ala	Pro	Gly	Thr	Glu	Pro	Ser	Pro	Ala	
240				260					265					270			
241	Pro	Leu	Gly	Gln	Pro	Arg	Ser	Tyr	Leu	Ala	Pro	Pro	Thr	Arg	Pro	Ala	
242			275					280					285				
243	Pro	Thr	Glu	Pro	Pro	Pro	Ser	Pro	Ser	Pro	Gln	Arg	Asn	Ser	Gly	Arg	
244		290					295					300					
245	Arg	Ala	Glu	Arg	Arg	Val	His	Pro	Asp	Leu	Ala	Ala	Gln	His	Ala	Ala	
246	305					310					315					320	
247	Ala	Gln	Pro	Asp	Ser	Ile	Thr	Ala	Ala	Thr	Thr	Gly	Gly	Arg	Arg	Arg	
248					325					330					335		
249	Lys	Arg	Ala	Ala	Pro	Asp	Leu	Asp	Ala	Thr	Gln	Lys	Ser	Leu	Arg	Pro	
250				340					345					350			
251	Ala	Ala	Lys	Gly	Pro	Lys	Val	Lys	Val	Lys	Pro	Gln	Lys	Pro	Lys		
252			355					360					365				
253	Ala	Thr	Lys	Pro	Pro	Lys	Val	Ser	Gln	Arg	Gly	Trp	Arg	His	Trp		
254		370					375					380					
255	Val	His	Ala	Leu	Thr	Arg	Ile	Asn	Leu	Gly	Leu	Ser	Pro	Asp	Glu	Lys	
256	385					390					395					400	
257	Tyr	Glu	Leu	Asp	Leu	His	Ala	Arg	Val	Arg	Arg	Asn	Pro	Arg	Gly	Ser	
258					405					410				415			
259	Tyr	Gln	Ile	Ala	Val	Val	Gly	Leu	Lys	Gly	Gly	Ala	Gly	Lys	Thr	Thr	
260				420					425					430			

VERIFICATION SUMMARY

DATE: 04/29/2003

PATENT APPLICATION: US/10/009,383

TIME: 08:08:00

Input Set : A:\07763-043001.TXT

Output Set: N:\CRF4\04292003\J009383.raw